

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-50254

(P2000-50254A)

(43) 公開日 平成12年2月18日 (2000.2.18)

(51) IntCl.

H 0 4 N 7/24

識別記号

F I

H 0 4 N 7/13

データベース (参考)

Z 5 C 0 5 9

審査請求 有 請求項の数24 OL (全 20 頁)

(21) 出願番号

特願平10-203916

(22) 出願日

平成10年7月17日 (1998.7.17)

(71) 出願人 597067574

ミツビシ・エレクトリック・インフォメイ
ション・テクノロジー・センター・アメリ
カ・インコーポレイテッドMITSUBISHI ELECTRIC
INFORMATION TECHNO
LOGY CENTER AMERIC
A, INC.アメリカ合衆国、マサチューセッツ州、ケ
ンブリッジ、ブロードウェイ 201

(74) 代理人 100057874

弁理士 曾我 道照 (外6名)

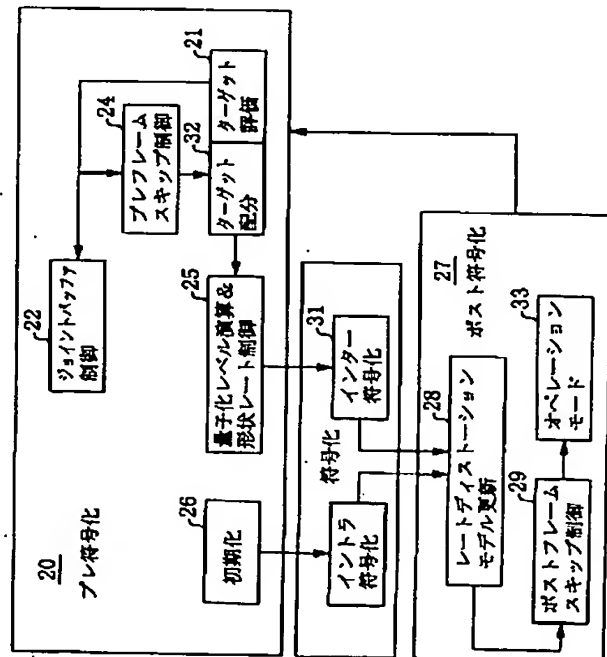
最終頁に続く

(54) 【発明の名称】 改良された適応性のあるビデオ符号化方法

(57) 【要約】

【課題】 MPEG-4エンコーダと他の符号化計画のために特に適した適応性のあるビデオ符号化の方法を提供する

【解決手段】 符号化ステップは、ビデオオブジェクトの各々のテクスチャ、動き及び形状情報を生成し、バッファに画像表示符号ビットを格納し、所定のリミット値に関してバッファの内容を制限し、各連続フレーム中の各ビデオオブジェクトを符号化するためのビットのターゲット数を評価し、少なくともより高いレートとより低いレートの内の一方で符号化するための可変レートを設定するステップを有する。



【特許請求の範囲】

【請求項1】 画像情報のフレームのシーケンスを適応性をもって符号化する方法において、可変入力ビットレートと実質的に一定の出力ビットレートとを有するバッファによって圧縮されたビデオ信号を送送チャネルへ供給するための、前記フレームの中の少なくともいくつかの複数のビデオオブジェクトを含み、変換係数を生成するために離散余弦変換を行うためのプロセッサと、可変レートで画像表示符号ビットを生成するために前記変換係数を量子化するための量子化装置とを含む符号化手段を用いて、一組のフレームの各々の前記ビデオオブジェクトの各々を符号化するステップであって、前記符号化ステップは前記ビデオオブジェクトの各々のテクスチャ、動き及び形状情報を生成し、前記バッファに前記画像表示符号ビットを格納し、前記フレームの連続フレーム中の前記ビデオオブジェクトのための前記符号化手段によって生成される符号ビットの量を増減するために、二次レートディストーションモデルに従って参照値について前記量子化装置によって利用される量子化パラメータを調整することによって所定のリミット値に関して前記バッファの内容を制限し、対応するオブジェクトプレーン中の対応するオブジェクトに関連する相対的動き、サイズ及び平方偏差パラメータの関数に従って、前記オブジェクトの間のビデオオブジェクトプレーン中の全てのオブジェクトについてビットのターゲット数を配分することにより、第1フレームに続いて所定の時間間隔にわたって発生するシーケンスの中で、各連続フレーム中の各ビデオオブジェクトを符号化するためのビットのターゲット数を評価し、画質を保ちながら前記バッファのオーバーフローを回避するために少なくともより高いレートとより低いレートの内的一方で符号化するための前記可変レートを符号化レートとして設定するステップを有することを特徴とする改良された適応性のあるビデオ符号化方法。

【請求項2】 前記関数がさらに、前記動き、サイズ及び平方偏差パラメータの各々に対する別個の重み係数を有することを特徴とする請求項1に記載の改良された適応性のあるビデオ符号化方法。

【請求項3】 前記平方偏差パラメータが、先行するオブジェクトプレーン中の対応する画素と比較して所定のオブジェクトプレーン中のビデオオブジェクトの各画素に対する平均絶対差値の演算から導出されることを特徴とする請求項1に記載の改良された適応性のあるビデオ符号化方法。

【請求項4】 前記方法がさらに、1つのフレーム中のすべてのビデオオブジェクトの符号化の最後にバッファビット容量と現在のバッファレベルとの間の差が所定のマージンより小さいときはいつでも、フレーム周期の間フレームの符号化をスキップするステップを有することを特徴とする請求項1に記載の改良された適応性のある

ビデオ符号化方法。

【請求項5】 相対的動き、サイズ及び平方偏差パラメータの前記関数が前記平均絶対差値の2乗に比例する変数を含むことを特徴とする請求項3に記載の改良された適応性のあるビデオ符号化方法。

【請求項6】 前記関数が、

$$T[i] = W_m \times MOT[i] + W_s \times SIZE[i] + W_v \times MAD^2[i]$$

であって、 $MOT[i]$ 、 $SIZE[i]$ 及び $MAD[i]$ は、動き、サイズ及び平均絶対差パラメータの相対比率を示し、 W_m 、 W_s 及び W_v は式

$$W_m + W_s + W_v = 1$$

を満たす重み係数であることを特徴とする請求項3に記載の改良された適応性のあるビデオ符号化方法。

【請求項7】 前記符号化レートが、前記より低いレートと前記より高いレートであるとき、 W_v がそれぞれより低い値とより高い値で選択されることを特徴とする請求項6に記載の改良された適応性のあるビデオ符号化方法。

【請求項8】 前記より低い符号化レートとして

$$W_m = 0.6, W_s = 0.4 \text{ 及び } W_v = 0$$

前記より高い符号化レートとして

$$W_m = 0.25, W_s = 0.25 \text{ 及び } W_v = 0.5$$

として前記重み係数が選択されることを特徴とする請求項7に記載の改良された適応性のあるビデオ符号化方法。

【請求項9】 前記より低い符号化レートとしての前記重み係数が $W_v = 0$ であって、前記より高い符号化レートとして W_v が W_m 又は W_s より大きいことを特徴とする請求項7に記載の改良された適応性のあるビデオ符号化方法。

【請求項10】 前記設定ステップの直前に先行するある時間間隔で、連続するスキップされたフレームの数をカウントすることによって前記可変レートの前記設定が決定されることを特徴とする請求項1に記載の改良された適応性のあるビデオ符号化方法。

【請求項11】 前記スキップされたフレーム数が所定の数より小さいとき、前記可変レートが前記より高いレートで設定されることを特徴とする請求項10に記載の改良された適応性のあるビデオ符号化方法。

【請求項12】 前記所定の数が2であることを特徴とする請求項11に記載の改良された適応性のあるビデオ符号化方法。

【請求項13】 前記スキップされたフレーム数が所定の数より小さいとき、前記可変レートが前記より高いレートで設定されることを特徴とする請求項4に記載の改良された適応性のあるビデオ符号化方法。

【請求項14】 画像情報のフレームのシーケンスを適応性をもって符号化する方法において、可変入力ビットレートと実質的に一定の出力ビットレートとを有するバッファによって圧縮されたビデオ信号を送送チャネルへ供給するための、前記フレームの中の少なくともいくつか

かが複数のビデオオブジェクトを含み、変換係数を生成するために離散余弦変換を行うためのプロセッサと、可変レートで画像表示符号ビットを生成するために前記変換係数を量子化するための量子化装置とを含む符号化手段を用いて、一組のフレームの各々の前記ビデオオブジェクトの各々を符号化するステップであって、前記符号化ステップは前記ビデオオブジェクトの各々のテクスチャ、動き及び形状情報を生成し、前記バッファに前記画像表示符号ビットを格納し、前記フレームの連続フレーム中の前記ビデオオブジェクトのための前記符号化回路によって生成される符号ビットの量を増減するために、二次レートディストーションモデルに従って参照値に関して前記量子化装置によって利用される量子化パラメータを調整することによって所定のリミット値に関して前記バッファの内容を制限し、対応するオブジェクトプレーン中の対応するオブジェクトに関連する相対的動き、サイズ及び平方偏差パラメータの関数に従って、前記オブジェクトの間のビデオオブジェクトプレーン中の全てのオブジェクトについてビットのターゲット数を配分することにより、第1フレームに続いて所定の時間間隔にわたって発生するシーケンスの中で、各連続フレーム中の各ビデオオブジェクトを符号化し、マスクに従って各オブジェクトについての前記形状情報を符号化し、所定の変換率に従って、符号化するための各前記オブジェクトの各マクロブロックのサイズ変換を行い、各前記マクロブロックの元のサイズを再構成し、前記マクロブロックの範囲内の各画素ブロックの変換誤差を決定し、エラー画素ブロックを識別するために前記変換誤差を所定の閾値と比較し、前記閾値を超過しないか最大変換率に到達するまで前記変換率を増やし、変換誤差と前記閾値との比較を再決定することを特徴とする改良された適応性のあるビデオ符号化方法。

【請求項15】 画質を保ちながら前記バッファのオーバーフローを回避するために少なくともより高いレートとより低いレートの内的一方で符号化するための前記可変レートを設定することを特徴とする請求項14に記載の改良された適応性のあるビデオ符号化方法。

【請求項16】 前記設定ステップの直前に先行するある時間間隔で、連続するスキップされたフレームの数をカウントすることによって前記可変レートの前記設定が決定されることを特徴とする請求項15に記載の改良された適応性のあるビデオ符号化方法。

【請求項17】 相対的動き、サイズ及び平方偏差パラメータの前記関数が前記平均絶対差値の2乗に比例する変数を含むことを特徴とする請求項16に記載の改良された適応性のあるビデオ符号化方法。

【請求項18】 前記関数が、

$$T[i] = W_m \times MOT[i] + W_s \times SIZE[i] + W_v \times MAD^2[i]$$

であって、 $MOT[i]$ 、 $SIZE[i]$ 及び $MAD[i]$ は動き、サイズ及び平均絶対差パラメータの相対比率を示し、 W_m 、 W_s 及び W_v は式

$$W_m + W_s + W_v = 1$$

を満たす重み係数であることを特徴とする請求項17に記載の改良された適応性のあるビデオ符号化方法。

【請求項19】 前記符号化レートが、前記より低いレートと前記より高いレートであるとき、 W_v がそれぞれより低い値とより高い値で選択されることを特徴とする請求項18に記載の改良された適応性のあるビデオ符号化方法。

【請求項20】 画像情報のフレームのシーケンスを適応性をもって符号化する方法において、可変入力ビットレートと実質的に一定の出力ビットレートとを有するバッファによって圧縮されたビデオ信号を送信チャネルへ供給するための、前記フレームの中の少なくともいくつかの複数のビデオオブジェクトを含み、変換係数を生成するために離散余弦変換を行うためのプロセッサと、可変レートで画像表示符号ビットを生成するために前記変換係数を量子化するための量子化装置とを含む符号化手段を用いて、一組のフレームの各々の前記ビデオオブジェクトの各々を符号化するステップであって、前記符号化ステップは前記ビデオオブジェクトの各々のテクスチャ、動き及び形状情報を生成し、前記バッファに前記画像表示符号ビットを格納し、前記フレームの連続フレーム中の前記ビデオオブジェクトのための前記符号化回路によって生成される符号ビットの量を増減するために、二次レートディストーションモデルに従って参照値に関して前記量子化装置によって利用される量子化パラメータを調整することによって所定のリミット値に関して前記バッファの内容を制限し、画質を保ちながら前記バッファのオーバーフローを回避するために少なくともより高いレートとより低いレートの内的一方で符号化するための前記可変レートを設定する所定の変換率に従って前記各オブジェクトの各マクロブロックの前記形状情報のサイズ変換を行い、前記マクロブロックの範囲内の各画素ブロックの変換誤差を決定し、エラー画素ブロックを識別するために前記変換誤差を所定のAlpha閾値と比較し、前記Alpha閾値を超過しないか最大変換率に到達するまで前記変換率を増やし、変換誤差と前記Alpha閾値との比較を再決定し、1つのフレーム中のすべてのビデオオブジェクトの符号化の最後にバッファビット容量と現在のバッファレベルとの間の差が所定のマージンより小さいときはいつでも、フレーム周期の間フレームの符号化をスキップし、前記設定ステップの直前に先行するある時間間隔で、連

統するスキップされたフレームの数をカウントすることによって前記可変レートの前記設定が決定され前記スキップされたフレーム数が所定の数より小さいとき、前記可変レートが前記より高いレートで設定され、符号化後に、前記より高いレート若しくは前記より低いレートの何れが有効であるかを判定し前記より低いレートが有効であれば前記Alpha閾値を増やし、前記より高いレートが後続の符号化間隔にとって有効であれば前記Alpha閾値を減らすステップを有することを特徴とする改良された適応性のあるビデオ符号化方法。

【請求項21】 前記Alpha閾値が初期値として実質的にゼロと最大値との中間の値に設定されることを特徴とする請求項20に記載の改良された適応性のあるビデオ符号化方法。

【請求項22】 前記Alpha閾値が最初の所定のレベルの増分の値で増加し、別の所定のレベルの減分の値で減少することを特徴とする請求項20に記載の改良された適応性のあるビデオ符号化方法。

【請求項23】 前記増分レベルが前記減分レベルより少ないことを特徴とする請求項21に記載の改良された適応性のあるビデオ符号化方法。

【請求項24】 対応するオブジェクトプレーン中の対応するオブジェクトに関連する相対的動き、サイズ及び平方偏差パラメータの関数に従って、前記オブジェクトの間のビデオオブジェクトプレーン中の全てのオブジェクトについてビットのターゲット数を配分することにより、第1フレームに続いて所定の時間間隔にわたって発生するシーケンスの中で、各連続フレーム中の各ビデオオブジェクトを符号化することを含むことを特徴とする請求項21に記載の改良された適応性のあるビデオ符号化方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、二次レートディストーションモデルに基づき複数のビデオオブジェクトに対してジョイントレート制御を用いてディジタル記録および/またはそのような信号の伝送を行うための、ビデオ信号を符号化する方法に関する。

【0002】特に、本発明は、ジョイントレート制御アルゴリズムを用いて複数のビデオオブジェクトの符号化を行うための、ビデオ符号化の方法に関するものである。このアルゴリズムは、MPEG-4ビデオ検証モデルV7.0 ISO/ICC JTC1/SC29/WG11、動画及び関連音声の符号化MPEG97/N1642. 1997年4月、Bristol, U.K. に解説されているVM7レート制御計画に基づくものである。

【0003】この方法は、ターゲット値の配分方法とレート制御プロセスにおいてオブジェクトの形状を考慮に入れるツールを導入する方法が変更されている。これらの変更は、ビデオオブジェクト間の均質的品質を増すこと及びより良いバッファの調整に貢献している。全体と

して、この方法は、バッファがよく調整され、ビットが適切に配分されるように複数のビデオオブジェクトを符号化する有効な手段を提供するものでありながら、それでいて空間品質と時間品質との間の必要な折衷を決定するための柔軟な方法である。

【0004】

【従来の技術】動画専門家グループ(MPEG)によって採用されたデジタルカラービデオ信号の帯域幅を圧縮する基本的方法是離散余弦変換(DCT)技術を利用するものである。さらに、MPEGのアプローチでは動き補償技術が用いられる。

【0005】画像の全フレームについての情報をほんの時々展開することによりMPEG規格では高いデータ圧縮率が達成される。この画像フレーム、すなわちイントラ符号化像は“I-フレーム”と呼ばれ、他のどんなフレームにも依存しない全フレーム情報を含んでいる。I-フレームの間に、いわゆるB-フレームとP-フレームがあり、これらのフレームには参照アンカーフレームに関連して発生する画像差のみが格納されている。

【0006】具体的には、ビデオシーケンスの各フレームは画素データのより小さなブロックに仕切られ、統計的に依存している空間領域ピクチャ成分(画素)を独立周波数領域DCT係数に変換するために各ブロックに離散余弦変換関数が施される。

【0007】すなわち、フレーム内符号化(I-フレーム)に従って符号化されたデータブロックは離散余弦係数のマトリックスから成る。それぞれの 8×8 乃至 16×16 の画素ブロックに離散余弦変換(DCT)が施され、符号化信号が生成される。これらの係数に適応量子化が行われ、次いでランレングス符号化と可変長符号化が施される。従って、送信データのそれぞれのブロックには 8×8 マトリックスの符号語より少ない符号語が含まれることがある。イントラフレーム符号化データのマクロブロックには、DCT係数に加えて、用いられている量子化のレベル、マクロブロックアドレスすなわち位置標識、及びマクロブロックタイプのような情報が含まれることになる。そして、この後者の情報は“ヘッダ”すなわち“オーバーヘッド”情報として参照される。

【0008】P又はBインターフレーム符号化に従って符号化されたデータブロックはまた離散余弦係数のマトリックスから成る。しかしこの場合、その係数は、予測された 8×8 画素マトリックスと実際の 8×8 画素マトリックスとの間の残余すなわち差を表す。これらの係数は量子化され、ランレングス符号化及び可変長符号化が施される。フレームシーケンスにおいて、I及びPフレームはアンカーフレームと呼ばれる。各Pフレームは、一番最後に発生するアンカーフレームから予測される。各Bフレームは、アンカーフレームの一方若しくはBフレームが間に配置されている2つのアンカーフレームから予測される。この予測符号化プロセスには、アンカーフレ

ームのどのブロックが現在符号化されている予測フレームのブロックにもっとも近似してマッチしているかを示す変位ベクトルの生成が伴う。アンカーフレームのマッチしたブロックの画素データは1画素ずつのベースで符号化されているフレームのブロックから減じられ、その残余が展開される。変換された残余とベクトルには予測フレームのための符号化データが含まれる。イントラフレーム符号化フレームと共に、このマクロブロックには量子化、アドレス及びタイプ情報が含まれる。

【0009】これらの結果は通常エネルギーが集中されているため、ブロック中の係数のいくつかはピクチャ情報の主要部分に含まれる。これらの係数は周知の方法で量子化され、係数のダイナミックレンジを有効に限定する。この結果は次いでランレングス及び可変長レングスで符号化され、伝送媒体に適用される。

【0010】ISO/IEC JTC1/SC29/WG11 MPEG96/N1469という名称で、1996年11月にMPEG-4ビデオVM編集に関する専門家班によってそのメンバーに配布された“MPEG-4ビデオ検証モデルバージョン5.0”に解説されている、最新の符号化検証モデル(VM)を実現するための最近の提案で、その内容は本明細書に参考として取り入れられているが、デービッドサーノフ(David Sarnoff)研究センターの代表者達は、“二次レートディストーションモデルを用いる新しいレート制御計画”を提案した。このMPEG-4ビデオ符号化フォーマットは(先行の計画の場合と同様)フレーム毎にエンコーダで可変ビットレートストリームを生成する。この可変ビットレートストリームは固定レートチャネルによって伝送されなければならないので、チャネルバッファを用いてビットストリームを円滑に出力する。バッファがオーバーフローまたはアンダーフローするのを回避するために、符号化プロセスのレート制御が必要とされる。

【0011】この最近の提案では、所定のセットのフレーム(ピクチャ)のために符号化プロセスが始まる前に、各フレームのためのターゲットビットレートが計算され、エンコーダの出力からの出力ビットレートが固定ビットレートに抑えられる一方、ピクチャ符号化の結果から生じるビットレートが画像フレームの内容によって(修正されないままになっている場合)比較的広い範囲にわたって変動可能であるという事実を助ける。この提案に従って、各フレームに関連する歪み測度はフレームの平均量子化スケールになるものと仮定されており、歪み測度の逆関数の二次関数としてそのレートディストーション関数がモデル化されている。実際の符号化プロセスが始まる前に、画像のターゲットビットレートは、符号化すべきフレーム数のみならず画像グループを符号化するために残されているビット数によっても評価される。この提案の提唱者達はピクチャレベルでの彼らの計画の実現について触れ、その計画をマクロブロックレベルにまで拡張する可能性についても述べている。

【0012】ブロック(マクロブロック)があるオブジェクトのエッジの境界を含むとき、DCT係数によって表されるように、変換後のそのブロックのエネルギーには比較的大きなDC係数(マトリックスの左上端コーナー)が含まれ、マトリックスの至る所にランダムにAC係数を配分したということも知られている。これに反して、非エッジブロックは、同様の大きなDC係数(左上端コーナー)及びそのブロックに関連する他の係数より実質的に大きいいくつかの(たとえば2つの)隣接AC係数によって通常特徴づけられる。この情報は空間領域の画像変化に関するものであり、連続フレーム(すなわち時間的差)の比較から得られる画像差情報と組み合わせられて1つのビデオオブジェクト(V0)を別のものと区別するために因数を利用できる。

【0013】図1(見本ビデオ場面)に示すように、2つ以上のビデオオブジェクト($V0_1$, $V0_2$, $V0_i$)が1つの画像フレームすなわちプレーン($V0P$)に含まれることがある。そして各連続フレームでは、ビデオオブジェクトの相対的位置が動きを表示しながら変化すると予想される。同時に、この動きがオブジェクトの画定を助ける。

【0014】MPEG-4VMに基づいて、システムのデコーダ端で付加的機能性を提供するためにコンテンツベースの操作と独立ビットストリーム符号化という追加目的が課された。このMPEG-4の目的は、MPEG-4エンコーダ内での形状情報の符号化のような付加オーバーヘッド情報の結果として各フレームに対する予測ターゲットビットレートのプロセスに関する付加的処理要件を複雑にし、押し付けるものである。個々のV0の同定に関する情報と同様MPEG-4VMの上述の特徴は、上述の参照マニュアルにより詳しい説明がある。

【0015】

【発明が解決しようとする課題】MPEG-4エンコーダと他の符号化計画のために特に適した適応性のあるビデオ符号化の方法を提供することが本発明の目的である。

【0016】各V0の相対的動き、サイズ、平方偏差及び形状を利用する改良されたビットレート制御システムを提供する際に個々のビデオオブジェクト(V0)を考慮に入れるMPEG-4VMに従って適応性のあるビデオ符号化を行うための利用可能な方法を提供することが本発明の更なる目的である。

【0017】

【課題を解決するための手段】本発明に従って、複数のビデオオブジェクトを符号化する際に一定のビットレート出力を行う方法を説明する。この実現には二次レートディストーションモデルが利用される。各オブジェクトはそれ自身のパラメータセットによって記述される。これらのパラメータを用いて、最初のフレームが符号化された後各オブジェクトに対する初期ターゲットビット評価が行われる。出力バッファの込み具合に基づいて、全ターゲット値が調整され、次いでフレームのオブジェク

トの活動を表すパラメータセットに比例して配分される。各オブジェクトに関連する動き、サイズ及び平方偏差パラメータから導出されるウェイトレシオを参照して活動が決定される。形状レート制御パラメータも呼び出される。新しい個々のターゲット値と二次モデルパラメータに基づいて、各ビデオオブジェクトに対する適切な量子化パラメータを計算することができる。この方法によって、待ち時間の短いビデオ符号化のためのターゲットビットレートの達成が保証される。

【0018】二次レートディストーションモデルに基づいて適切なビットレート制御システムを提供するために、フレーム全体に関してではなくビデオオブジェクト(V0)に関して制御情報を一緒に適用できることがわかっている。

【0019】

【発明の実施の形態】実施の形態1. ジョイントビットレート制御を行うための方法はプレ符号化段階とポスト符号化段階に分けることができる。図2に示すように、プレ符号化段階20には、(i)ターゲットビット評価21、(ii)ジョイントバッファ制御22、(iii)プレフレームスキップ制御24、及び(iv)量子化レベル演算25が含まれる。ポスト符号化段階27には、(i)レートディストーションモデルの更新28と(ii)ポストフレームスキップ制御29が含まれる。この計画の重要な側面は、そのブロック構造からは明らかではないが、ほとんどのブロックによって、全てのビデオオブジェクト(V0)について前オペレーションが完全であるように要求されるという点である。たとえば、符号化される次のV0のインター符号化31は、先行するV0に対するすべての量子化レベルが計算されてしまってから初めて計算される。この実施の形態では、すべてのV0が同一のフレームレートで符号化される。しかし、現在の実施の形態の多くは、各V0に対する異なるフレームレートへの移動を予測するものである。しかし、より完全なバッファ制御が要求される。

【0020】適応性のあるビデオエンコーダは、図2に

例示する方法に従うようになっている。デジタル画像ソース10は、フレームベース又はビデオオブジェクト(V0)ベースの画像情報をビデオ信号エンコーダ12にデジタル化された形で提供する。この画像は区画され、空間的に重複しない画素データブロックになる。8×8画素若しくは16×16画素のブロックサイズを用いてもよい。それから各区画ブロックは処理される。

【0021】動き評価部14は、フレームメモリ(以前に再構成されたVOP16)に格納された時間的に近い参照フレームを参考にして入力ブロック動きベクトルを評価するために用いられる。この参照フレームは、元々ある未処理フレーム若しくは以前に符号化されたフレームであってもよい。MPEG規格に記載されているような両方向の動き評価を適用してもよい。

【0022】動き補償ブロック11、テクスチャ符号化ブロック13、形状コーディングブロック15、一定出力ビットレートバッファ17及び多重送信MSDL19、これらすべてはMPEG-4の参考文献に記載されているように構成され、提供される。これに加えて、機能を果たすために(図2に関連して説明したような)レート制御システム18が提供される。

【0023】再び図2を参照するが、プレ符号化段階20にはさらに初期化セクション26(以下に示す表1を参照)が含まれる。

【0024】A. 初期化セクション26

このセクションでは、各V0がそれ自身の個別情報を担うことができるように、ほとんどのレート制御変数(たとえば、一次及び二次の複雑さとMADすなわち平均絶対差情報)はベクトルにまで拡張される。セグメントに対する残りビット数とバッファドレインのレートは変更されない。表1はこの方法の説明のために用いられる表記法を要約するものである。

【0025】

【表1】

表 1

二次レートディストーションモデルに基づくジョイントレート制御に用いられる表記法

変数	説 明
Buff_drain	ピクチャ当りバッファから取り除かれるビット数
Mad[i]	動き補償後の現在の VO に対する平均絶対差
X1[i], X2[i]	一次及び二次複雑さ測定値
Q[i]	i 番目の VO に対する量子化パラメータ
N_skip_post	ポストフレームスキップによりスキップすべきフレーム数
N_skip_pre	プレフレームスキップによりスキップすべきフレーム数
N_skip	スキップすべきフレーム総数
N_btwn	符号化されたフレーム間のフレーム数
B_left	シーケンス符号化のために残されたビット数
T_texture[i]	i 番目の VO に対するテクスチャビットカウント
T_texture	総テクスチャビットカウント (全ての VO)
T[i]	テクスチャ、形状、動き及びヘッダビットを含む i 番目の VO に対するビットカウント
T	テクスチャ、形状、動き及びヘッダビット (全ての VO) を含む総ビットカウント
H[i]	形状と動きを含むヘッダビットカウント
H	総ヘッダビットカウント (全ての VO)
Buff_size	バッファサイズ
Buff_level	現在のバッファの込み具合

【0026】B. ポスト符号化段階27

符号化段階30の後に、レートディストーションモデルのためのパラメータを探さなければならない。多くのVOに対して、エンコーダレートディストーション関数は次のようにモデル化される。

$$T_{\text{texture}} = \sum_{\text{all vo}} (X1_i \cdot \text{MAD}_i + X2_i \cdot \text{MAD}_i)$$

$$\left(\begin{matrix} Q_i & Q_i^2 \end{matrix} \right)$$

上記の式から、全てのVOに対して別個にモデルパラメータ $X1_i$ と $X2_i$ を計算できる。上記の式において、ターゲット値、 T_{texture} は複数の T_{texture_i} に分解されるが、これはi番目のVOのみから成るテクスチャ成分を符号化するために使われるビット総数に対応するものである。

【0027】図2では、ポスト符号化段階27の次のステップはポストフレームスキップ制御関数29である。この時点でバッファ17は更新される。スキップマージン γ に対して現在のバッファレベルをチェックすることによってオーバーフローが回避される。もし現在のバッファレベルが指示されたマージンより上にあれば、フレームは継続的にスキップされる。すなわち、特定の基準が満たされるまで $N_{\text{skip_post}}$ は1つずつ増える。

【0028】このポストフレームスキップ制御は基準 $\text{Buff_level} - N_{\text{skip_post}} \cdot \text{Buff_drain} < (1 - \gamma) \cdot \text{Buff_size}$

が満たされるまで1つずつ増える。

【0029】 γ は、推奨構成では0.2に等しくなるように選ばれる。上記の式の条件が満たされ、 $N_{\text{skip_post}}$ が見いだされた後、 $N_{\text{skip_pre}}$ の値が $N_{\text{skip_post}}$ に加

えられる。 $N_{\text{skip_pre}}$ の確定については後程説明する。この最終値、 $N_{\text{skip}} = N_{\text{skip_pre}} + N_{\text{skip_post}}$ は、スキップされる全フレームに等しい。新しい時刻を確定するのはこの値である。ポストフレームスキップ制御関数が発生した後でしかこの時刻は更新できないことに留意されたい。

【0030】最初の構成の次のプレ符号化段階を続行しながら、セグメントの利用可能なビット数と前回の対応するVOで使われたビット数に基づいて初期ターゲットビットレートを計算する。最低限の品質が保証されるように、フレームベースのシミュレーションに対する同様の下限を用いる。

$$T[i] = \text{Max}(B_{\text{left}} / (30 \cdot \text{numVOs}), B_{\text{left}} / (\text{numVOleft}[i] \cdot \text{numVOs}))$$

$$T[i] = T[i] \cdot (1 - \alpha) + B_{\text{past}}[i] \cdot \alpha$$

【0031】重み付け α は過去のフレームからの寄与を表し、この実施の形態では0.2にセットされている。

【0032】いったん初期ターゲット値がセットされると、 $T = T \cdot (2a + b) / (b + 2a)$ (但し $a = \text{Buff_level}$, $b = \text{Buff_size} - \text{Buff_level}$) に従って、バッファ18の込み具合に基づく調整が行われる。このターゲットレートは個々のターゲットレート全ての総和を表すことに留意されたい。さらに現在及び将来のバッファ18レベルに関してはターゲット値の予想される影響に基づいて変更が行われる。 δ で安全マージンを示し、以下の式によりこのターゲット値を増大させる。

$$T_{\text{inc}} = \text{Buff_drain} - T - \text{Buff_level} + \delta \cdot \text{Buff_size}$$

$$\text{但し、} \text{Buff_level} - \text{Buff_drain} + T < \delta \cdot \text{Buff_size}$$

これに反して、以下の式によりターゲット値を減少させ

る。

$$T_{dec} = \text{Buff_level} + T - (1 - \delta) \cdot \text{Buff_size}$$

但し、 $\text{Buff_level} + T > (1 - \delta) \cdot \text{Buff_size}$

【0033】上記のオペレーションはジョイントバッファ制御の一部である。この例示の実施の形態では、 δ は0.1にセットされている。

【0034】次のステップは、一貫性のある品質を異なるオブジェクトにわたって維持するようにビットを再配分することである。これを達成するために、オブジェクトサイズとオブジェクトが経験する活動量を先に符号化したオブジェクトのヘッダ情報から取得する。しかし、このターゲット値を配分する前に、以前のフレームのヘッダに使用されたビット総数がこのビットカウントを超過するかどうかを判定するチェックが行われる。その差、 $s = T - H$ が全てのV0のテクスチャを符号化するために利用可能なビット数の近似値を表す。 $s < 0$ ならば、各V0を均等に符号化するのに十分なビットがないかもしれない。この場合、全てのターゲット値は負になる。後程説明するが、このことは下限制約条件を量子化パラメータに強いることになり、それによってテクスチャに費やされるビット総数を制限することになる。それに加えて、 $s < 0$ の場合、プレフレームスキップ制御24が呼び出される。ポスト符号化段階27の後時刻が更新されるだけなのでこのフレームスキップ制御ブロックは次の時刻更新に向けての補正として機能する。プレフレームスキップ制御24が呼び出されると、 N_skip_pre のノンゼロ値が決定される。この値は以下の式に従って決定される。

```
while(s < 0)
{
  increment N_skip_pre
  s = s + Buff_drain
}
```

【0035】ターゲット値を負にし余分のフレームをス

```
T_texture_i = Max( Buff_drain
                    -----, T_texture_i
                    { 3 * MAD_i * numVOs
}
```

【0039】ターゲット値が負の場合、導出された量子化パラメータはLB_QUANTによって下限が設けられ、負でない場合には1~31の通常のクリップ値が用いられる。このパラメータを使用することによって比較的少量のビットがテクスチャの符号化に行くことが保証される。LB_QUANTの値は25より大きくなるように選ばなければならない。代わりの方法として、アルファブレーションの解像度を符号化の前に減少させることによって形状コーディン

表2

複数のビデオオブジェクトの符号化のための実験グループ

ID	シーケンス	ビットレート	フレームレート	フォーマット
1	アキヨ, コンテナ	10	7.5	QCIF
2	アキヨ, コンテナ	24	10	QCIF
3	ニュース	48	7.5	CIF
4	沿岸警備隊員	48	10	QCIF
5	沿岸警備隊員	112	15	CIF

【0042】複数のV0の符号化の際、3つのパラメー

キップするこの組合せによりレート制御アルゴリズムが次のターゲット値をより良く評価することが可能になり、一定のオブジェクト品質が供給される。

【0036】 $s > 0$ の場合には、先に符号化したフレームのヘッダ情報に比例して以下のように全ターゲット値Tが配分される。

$$T_i = H_i \cdot \left(1 + \frac{s}{H} \right)$$

【0037】各V0のターゲット値を決めると、次のタスクは所望のレートに対応する個々の歪み測度を決定することである。このプロセスを各V0について別個に処理しMADに関して標準化すると古典的二次方程式を解くことになる。

$$ax^2 + bx + c = 0,$$

但し $a = X2_i$

$b = X1_i$

$$c = T_texture_i / MAD_i$$

【0038】 $T_texture_i$ がヘッダビットを総ビット数から差し引いた結果であることに留意すれば、低ビットレートの符号化を行うとき小さなターゲット値を取得することが可能である。この困難を克服するために、以下の式に従ってターゲット値の下限を設定する。

グに費やすビット総数を減らしてもよい。

【0040】

【実施例】以下の表には上記セクションで説明したアルゴリズムの実験グループが要約されている。15という初期量子化パラメータをI-フレームに対して選んだが、その後この量子化パラメータは自動的に決定された。

【0041】

【表2】

タ、すなわち、形状、動き及びテクスチャが符号化され

る。エンコーダソフトウェアによって、ユーザーは組合せモード若しくは分離モードでパラメータを符号化することができる。本明細書に示すシミュレーション結果は分離パラメータ符号化を用いたものである。表3に各VO

に対するPSNRの詳細(Y-成分のみ)を示すとともに達成された実際のビットレートを報告する。

【0043】

【表3】

表3

コア実験のシミュレーション結果

シーケンス	ターゲットレート	フレームレート	フォーマット	Y-psnr V00	Y-psnr V01	Y-psnr V02	Y-psnr V03	Y-psnr V04	Y-psnr V05	実際のレート
アキヨ	10	7.5	QCIF	35.79	27.23	N/A	N/A	N/A	N/A	10.03
コンテナ	10	7.5	QCIF	31.96	23.76	19.98	25.56	29.60	18.87	9.90
アキヨ	24	10	QCIF	40.42	29.82	N/A	N/A	N/A	N/A	23.90
コンテナ	24	10	QCIF	31.54	22.73	19.75	25.37	29.08	18.62	23.78
ニュース	48	7.5	CIF	35.87	28.50	27.83	24.68	N/A	N/A	47.08
沿岸警備隊員	48	10	QCIF	29.09	23.15	25.96	26.40	N/A	N/A	47.82
沿岸警備隊員	112	15	CIF	27.13	22.74	24.24	26.64	N/A	N/A	111.94

【0044】これらの結果から、満足すべき画質をもつターゲットビットレートを達成することが可能であることは明らかである。しかし、待ち時間の短い利用のためにはバッファのオーバーフローを回避しなければならない。複数のビデオオブジェクトを符号化するためには、指定しなければならない多量のオーバーヘッド情報が必要である。このために、各オブジェクトのテクスチャを符号化するためのビットがエンコーダにほとんど残されていないということがよく起こる。その結果、エンコーダは、適切なバッファレベルを維持するためにフレームのスキップおよび/または量子化レベルの増加を余儀なくされる。これに加えて、符号化フレームの数が指定される。各シーケンスについて、300フレーム(10秒)が符号化された。

【0045】二次レートディストーションモデルに基づく1つの場面の複数のビデオオブジェクトの符号化手段を提示した。この計画は、フレームベースの符号化シミュレーションを行うためにすでに証明された方法を強化したものである。フレームスキップ制御が呼び出され、バッファを助けてそれが混みすぎないようにする。バッファがオーバーフローした事例は、非常に大量のビットがオーバーヘッドに費やされたことを示すものである。これを回避するために、オーバーヘッド情報(たとえば、形状に費やされるビット)量を減少してもよい。

【0046】本発明に従って、図3に示すような、ターゲット値配分32に関する改変を伴う基本的アプローチ、オペレーションモード33、形状に関連するレート制御25'及びポストフレームスキップ29'が用いられる。2つのオペレーションモードが用いられる。ターゲット値の配分は各オブジェクトのサイズ、動き及び平方偏差(すなわちMAD)に基づく。各配分に対する重みはオペレーションモードに依存する。最初のモードは低いビットレート(LowMode)に関するもので、次のモードは高いビットレート(HighMode)に関するものである。この新しいターゲット値の配分によって主観的品質に重要な改良点が生み

出される。フレームスキップ制御の改変によってレジスタのオーバーフローが回避される。

【0047】ジョイントレート制御を行うための方法はプレ符号化段階とポスト符号化段階に分けることができる。上述のように、プレ符号化段階20には、(i)ターゲットビット評価21、(ii)ジョイントバッファ制御22、(iii)プレフレームスキップ制御24、及び(iv)量子化レベル演算25'が含まれる。ターゲットビット評価21は以下に説明する修正ターゲット値配分関数32にも関連する。量子化レベル演算25'は以下に説明する形状レート制御関数にも関連する。ポスト符号化段階27には、(i)レートディストーションモデルの更新28、(ii)ポストフレームスキップ制御29'とオペレーションモード関数33が含まれる。図3は本発明のレート制御プロセスを例示し、加算ターゲット値配分32、オペレーションモード33、形状関連レート制御25'及び修正ポストフレームスキップ制御29'を含む本発明に関連する付加的特徴が含まれる。

【0048】場面の全てのオブジェクトについてターゲット値が求められ、同一のフレームレートで全てのビデオオブジェクトが符号化された。そして、1つのフレームの総ビット数は前回の対応するオブジェクトのヘッダビット量に比例して配分された。本発明の場合には、各オブジェクトの相対的動き、サイズ及び平方偏差すなわち“MAD”を考慮に入れる関数に比例してビットが配分される。各VOPの各々の特別なビデオオブジェクト(これは補償された動きであるが)に関連するMAD(平均絶対差)は、レート制御を行う目的で平方偏差の適切な指標となるように決定される。推奨構成では、 $MAD[i]$ 因数は $MAD^2[i]$ となるように選択される。総ターゲット値Tとして、全てのVOのビット総数は次のように表すことができる。 $T[i] = W_m \times MOT[i] + W_s \times SIZE[i] + W_v \times MAD^2[i]$ 、但し $MOT[i]$ 、 $SIZE[i]$ 及び $MAD^2[i]$ はそれぞれ、動き、サイズ、及び平均絶対差パラメータの相対比率を示し、 W_m 、 W_s 及び W_v は次式を満たす重みである。

$$W_m + W_s + W_v = 1$$

【0049】オペレーションモード

ポスト符号化関数27にはオペレーションモード関数33が含まれる。具体的には、2つの異なるオペレーションモード(低いビットレートでの符号化モードと高いビットレートでの符号化モード)が提供される。

【0050】高いビットレートで符号化するとき、ビットが自由に利用できると各VOへのターゲット値割り当ての際プロセスを柔軟にすることが可能になる。これらの環境の下では、均質の品質を各VOに課することが合理的である。従って、MADパラメータを含むことがターゲット値配分にとって重要であり、もっとも高いウエートを占める事項となる。これに反して、ビットが自由に利用できることが制限されていると様々なVOの間で均質の品質を達成することが非常に困難になる。また、低いビットレートという制約条件の下では、背景に費やすビットを少なくし前景により多くのビットを費やすことが望ましい。その場合、MADパラメータの重要性は小さくなり、動きパラメータの重要性が大きくなる。上記の論拠及び実験的試行錯誤に基づいて、推奨の重みはLowModelについては $W_m=0.6$ 、 $W_s=0.4$ 、 $W_v=0.0$ 、HighModelについては $W_m=0.25$ 、 $W_s=0.25$ 、 $W_v=0.5$ となる。

【0051】各フレーム内で品質を調整するほかにも、時間的品質の調整すなわち、フレームスキップを最小限にすることも重要である。HighModeでは、ビットをたっぷり自由に使用できるのでこれを行うことは非常に簡単である。しかし、LowModeでは、フレームスキップがずっと頻繁に発生する。実際、スキップされるフレーム数はそのプロセスを作動させるべきモードをよく示すものである。これは次のように表現される。

```
if(total_frame_skipped)SKIP_TH)
    Operate in LowMode
else
    Operate in HighMode
```

この実施例では、スキップ閾値(SKIP_TH)は、2に設定された。

【0052】オペレーションモードを獲得する決定プロセスは、時間的問題の解決に対する制約と見られることもできる。システムがLowModeの場合、エンコードはいくつかの指定された数のフレームをスキップした。空間的品質と時間的品質との間の妥当な折衷的解決を得るために、LowModeでは量子化パラメータの計算値の下限が課せられる。

【0053】ポストフレームスキップ制御29'の修正関数は現在のバッファ17占有率を決定し、将来のビデオオブジェクトの符号化がバッファ17をオーバーフローさせないことを保証するものである。 N_{skip_post} の正の値は次の条件を満たすように定められた。

$$Buff_level - N_{skip_post} \cdot Buff_drain < (1 - \gamma) \cdot Buff_size$$

【0054】本発明の実施例では、伝送する必要があるようなビット総数をよりよく予想するために前回のフレームからの情報が利用される。その新しい条件とは以下の通りである。

$$Buff_level + B_last - (N_{skip_post} + 1) \cdot Buff_drain < (1 - \gamma) \cdot Buff_size$$

但し、 B_last は、前回のフレーム又はビデオオブジェクトのセットを符号化するのに費やされた総ビット数を示す。この方法で、バッファ17は前回の符号化時に費やされた同量のビットをすぐに受け入れる。超過ビットは安全マージン中へ吸収され、その後、オーバーフローの発生を回避する。前回と同様、ガンマパラメータ、すなわちスキップマージンは0.2となるように選ばれる。

【0055】形状関連レート制御

特定のオブジェクトを画定するバイナリ形状情報(すなわちバイナリアルファプレーン)は、それがオブジェクトの一部である場合画素値を"255"まで設定したり、それがオブジェクトの外側にある場合画素値を"0"に設定するマスクにすぎない。MPEG-4ビデオ検証モデルのバージョン7.0によれば、形状情報のレート制御とレート低下はアルファプレーンのサイズ変換を通じて達成することができる。可能な変換率(CR)は1、1/2、または1/4である。換言すれば、 16×16 のマクロブロック(MB)はダウンコンバートして 8×8 若しくは 4×4 のブロックにすることができる。オブジェクトの相対的形状情報を含む各マクロブロックは、符号化するためにダウンコンバートし、次いで元のサイズで再構成することができる。全ての 4×4 画素ブロック(PB)について変換誤差が計算される。この変換誤差は、元のPBの一つの画素値とその再構成されたPBとの間の絶対差の和として定義される。この変換誤差が、"アルファ閾値"の16倍(すなわち、 $16 \times AlphaTH$)より大きい場合には、このPBは"エラーPB"と呼ばれる。1つの"エラーPB"がマクロブロック中にあれば、マクロブロックの変換率(CR)は最大値が1になるまで増加される。

【0056】形状コーディングについての上記の説明から、 $AlphaTH$ の値が各VOの形状情報に費やされるビット数に対してかなりの影響力を持っていることが明らかである。 $AlphaTH$ の値の選択及び2つのオペレーションモード(LowModeとHighMode)に基づいて形状情報を制御するための本発明の方法を説明する。

【0057】 $AlphaTH$ は初期値として $AlphaINI$ という値に設定されるものとする。I-フレームと第1P-フレームの処理中、この初期値はそのフレーム中の全てのオブジェクトの形状を符号化するために利用される。この符号化段階後、レート制御アルゴリズムによってオペレーションモードが決定される。オペレーションモードがLowModeに決定された場合、システムは現在の $AlphaTH$ を $AlphaINC$ に増大させる。オペレーションモードがHighModeに決定された場合、システムは現在の $AlphaTH$ を $AlphaDEC$

に減少させる。AlphaTHの最大値及び最小値はそれぞれAlphaMAX及び0である。この形状レート制御アルゴリズムは図4に要約されている。横軸は時間を示す。この軸に沿ってオペレーションモードを見分けるマーク(H - HighMode、L- LowMode)がある。縦軸は各符号化時刻における対応するAlphaTHを示す。この例では、AlphaMAXは16に設定されており、初期値AlphaINI=8(最大値の2分の1)である。また、この例ではAlphaDEC=5、AlphaINC=3である。もっとも一般的な場合、AlphaINCとAlphaDECは定数である必要はなく、むしろ現在のAlphaTHの関数(たとえば、ゼロに近いときにはより大きなステップ幅で、MAXに近いときにはより小さなステップ幅で)であればよい。推奨の構成では、AlphaMAX=12、AlphaINC=AlphaDEC=4そしてAlphaINI=0である。

【0058】図4では各符号化時刻に行われたアクションが強調されているということに留意されたい。各符号化時刻は均等な間隔で配置されている。実際のシミュレーションでは、前回のポスト符号化段階でスキップされたフレーム総数がskip_THの選択値より大きくなった後にLowModeが作動するだけなので、時間符号化時刻が非均一なものになっている。

表4

複数のビデオオブジェクトの符号化のための高いビットレート実験グループ

ID	シーケンス	ビットレート	フレームレート	フォーマット
6	アキヨ, コンテナ	48	10	QCIF
7	ニュース	192	15	CIF
8	沿岸警備隊員	384	30	CIF

【0062】表5に、低-ビットレート条件下での各VOに対する平均PSNR(ピーク信号対雑音比)値を示す。符号化フレーム数、各ビデオオブジェクトの範囲内の平均量子化スケール及び達成された実際のビットレートも示されている。高いビットレートシミュレーションについての同じ情報を表6に示す。図6～16の各テストシーケンスの

【0059】オペレーションモードに基づくAlphaTHのこの適応性のある選択は、非常に低いビットレートで十分な品質を維持しながら、形状に要求されるビット数を減少させる際極めて有効である。高いビットレートで、あるいはLowModeが頻繁には起こらないシミュレーションでは、低いAlphaTH値を使って形状情報を符号化することができ、予想されるように非常に高い品質のオブジェクト境界が結果として生じる。この方法は追加的機能性をオペレーションモードに与え、追加的テクスチャビットを解放しおよび/または適切なパッファ占有率を維持することによって時間的、空間的符号化の解決を調整するその努力を補完するものである。

【0060】

【実施例】上記の表2には、低いビットレートシミュレーションのテスト条件が示されている。表4には、高いビットレートシミュレーションのテスト条件が示されている。それぞれの表で、15という初期量子化パラメータをI-フレームに対して選んだが、その後この量子化パラメータは自動的に決定された。

【0061】

【表4】

パッファ占有率の線グラフに、低いビットレートと高いビットレートの条件下でこの方法によって示される例外的制御を例示する。

【0063】

【表5】

表 5

低いビットレート条件下の実験グループのシミュレーション結果

シーケンス ID	ビデオオブジェクト	平均量子スケール	PSNR (符号化 VO)	符号化フレーム	PSNR (全フレーム)	実際のビットレート
アキヨ-1 コンテナ-1	VO0	29.6	34.20	64/75	30.26	9.86
	VO0	30.4	31.13	29/75	26.16	
	VO1	30.0	23.24			
	VO2	27.6	19.91			
	VO3	27.6	24.13			
	VO4	29.6	28.60			
	VO5	27.6	18.48			
アキヨ-2	VO0	20.3	36.71	100/100	33.48	23.87
	VO1	9.7	31.73			
コンテナ-2	VO0	29.7	31.21	77/100	27.21	23.92
	VO1	28.7	23.08			
	VO2	10.7	26.64			
	VO3	12.9	26.63			
	VO4	29.6	28.65			
	VO5	6.8	27.96			
ニュース-3	VO0	29.3	34.66	62/75	26.16	47.68
	VO1	28.8	28.38			
	VO2	29.3	27.44			
	VO3	6.3	29.59			
沿岸警備隊員-4	VO0	26.4	28.75	93/100	25.11	47.69
	VO1	17.5	25.10			
	VO2	12.4	27.52			
	VO3	22.2	26.70			
沿岸警備隊員-6	VO0	30.6	27.14	76/150	23.05	111.27
	VO1	29.9	22.99			
	VO2	28.4	24.77			
	VO3	30.6	26.53			

【0064】

【表 6】

表 6

高いビットレート条件下の実験グループのシミュレーション結果

シーケンス ID	ビデオオブジェクト	平均量子スケール	PSNR (符号化 VO)	符号化フレーム	PSNR (全フレーム)	実際のビットレート
アキヨ-6	VO0	4.2	42.68	147/150	37.90	63.48
	VO1	4.9	36.27			
コンテナ-6	VO0	25.8	32.18	150/150	29.52	63.28
	VO1	15.7	28.05			
	VO2	4.2	32.04			
	VO3	4.3	36.06			
	VO4	26.9	30.03			
	VO5	2.11	39.23			
ニュース-7	VO0	3.7	43.43	148/150	33.50	254.23
	VO1	7.1	36.06			
	VO2	8.5	33.54			
	VO3	4.7	45.68			
沿岸警備隊員-8	VO0	23.5	28.36	299/300	28.23	383.78
	VO1	15.6	27.09			
	VO2	11.2	29.40			
	VO3	18.1	29.08			

【0065】上述の発明に従ってターゲット値配分に関する改良が行われた。また、形状レート制御のメカニズムが実現された。各実験グループのシミュレーションで優れた改良点が表示されている。提案されたジョイントレート制御計画のハイライトとしては、良好な主観的品質、優れたバッファ調整、VO間の均質な品質、形状とテクスチャの符号化のジョイント制御及び空間品質と時間品質を折衷させるための柔軟な構成が挙げられる。

【0066】値配分に対する改変は1つのオブジェクトの範囲内で平方偏差をよりよくモデル化する機能を果たしている。この平方偏差は符号化に必要なビット総数を示すために伝統的に用いられてきたもので、オブジ

ェクト間の歪みはより一貫性のあるものになる。

【0067】オペレーションモードに基づくAlphaTHの適応性のある選択は、非常に低いビットレートで十分な品質を維持しながら同時に形状を表すためのビット数を減少させるのに極めて有効である。高いビットレートで、あるいはLowModeが頻繁には起こらないシミュレーションでは、低いAlphaTH値を使って形状情報を符号化することができ、非常に高い品質のオブジェクト境界が結果として生じる。この方法は追加的機能性をオペレーションモードに与え、追加的テクスチャビットを解放しおよび/または適切なバッファ占有率を維持することによって時間的、空間的符号化の解決を調整するその努力

を補完するものである。

【0068】

【発明の効果】全体として、この方法によって、待ち時間の短い低いビットレートを適用して複数のビデオオブジェクトを符号化するという観点からMPEG-4規格の機能性を調整することが可能となる。より高いビットレートへこれを適用して測定できることも実証されている。

【0069】本発明を推奨実施例について説明してきたが、以下の請求の範囲に記載されている本発明の範囲から逸脱せずにこの実施例の詳細について様々な改変を行うことができる。

【図面の簡単な説明】

【図1】 VOの各々が左から右へと場面の中で時間につれて動く2つのビデオオブジェクト(VO_1 と VO_2)と背景画像を持つ3つの連続画像フレームの図式的絵による表示を示す図である。

【図2】 ステップ間の相互関係におけるステップを例示するブロック図である。

【図3】 ステップ間の相互関係と共に本発明の方法におけるステップを例示するブロック図である。

【図4】 本発明を実現するために改変されたMPEG-4エンコーダのブロック図である。

【図5】 本発明のオペレーションモード("H"又は"L")に基づく形状レート制御決定の時間的シーケンスを例示する、時間関数としてのパラメータ"AlphaTH"を示す図である。

【図6】 表2で確認できるような本発明に従うロー及

びハイビットレートで符号化されたID1アキヨの画像を表す一連のビデオ信号シーケンスに対するフレームのバッファ占有(ビット)率を示す線グラフである。

【図7】 表2のID1コンテナの画像のバッファ占有率を示すの線グラフである。

【図8】 表2のID2アキヨの画像のバッファ占有率を示すの線グラフである。

【図9】 表2のID2コンテナの画像のバッファ占有率を示すの線グラフである。

【図10】 表2のID3ニュースの画像のバッファ占有率を示すの線グラフである。

【図11】 表2のID4沿岸警備隊員の画像のバッファ占有率を示すの線グラフである。

【図12】 表2のID5沿岸警備隊員の画像のバッファ占有率を示すの線グラフである。

【図13】 表4のID6アキヨの画像のバッファ占有率を示すの線グラフである。

【図14】 表4のID6コンテナの画像のバッファ占有率を示すの線グラフである。

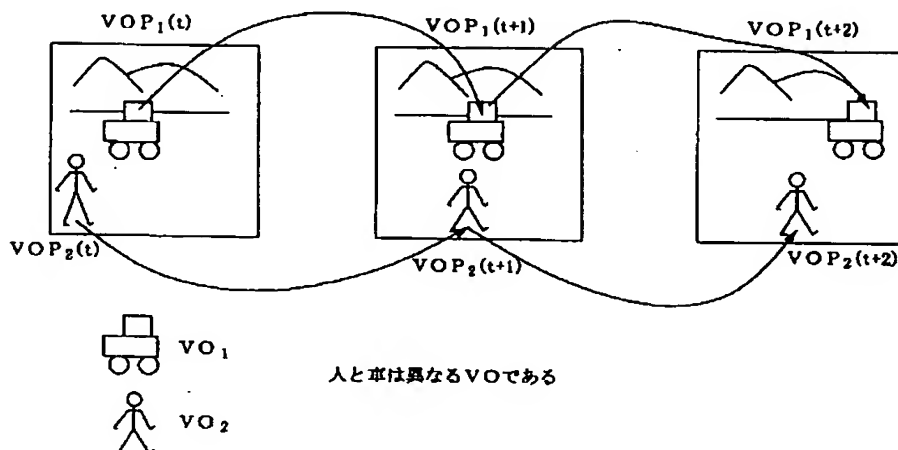
【図15】 表4のID7ニュースの画像のバッファ占有率を示すの線グラフである。

【図16】 表4のID8沿岸警備隊員の画像のバッファ占有率を示すの線グラフである。

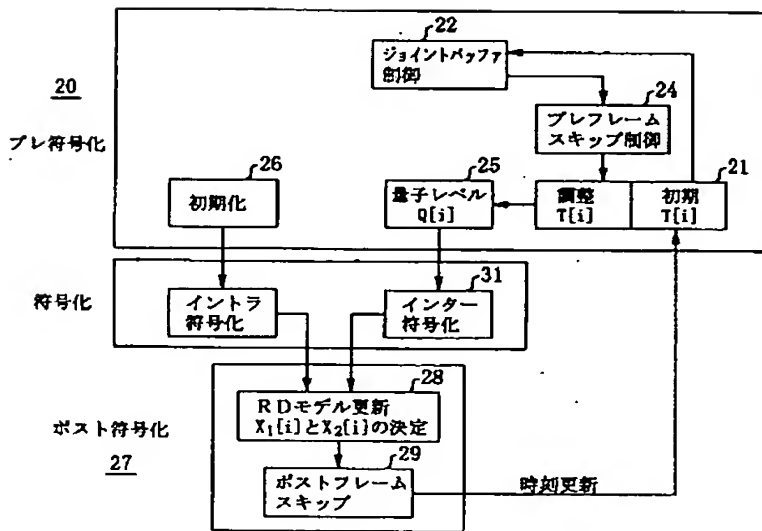
【符号の説明】

VO ビデオオブジェクト、25 量子化レベル演算(量子化装置)、20 プレ符号化段階(符号化手段)、27 ポスト符号化段階(符号化手段)。

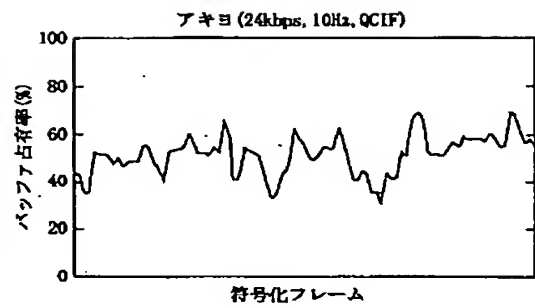
【図1】



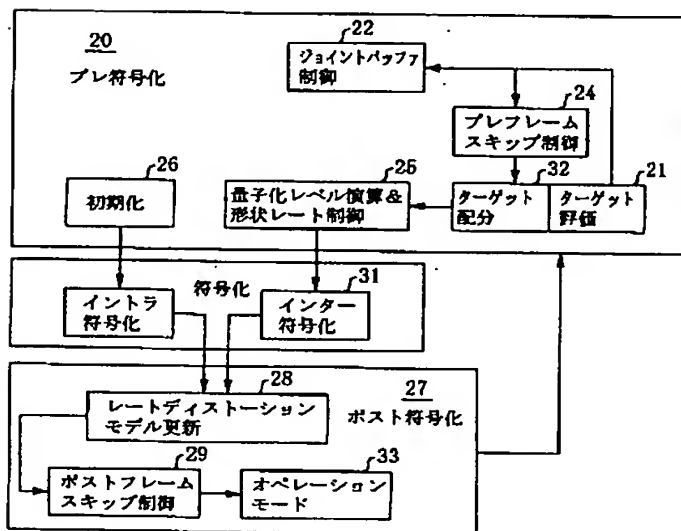
【図2】



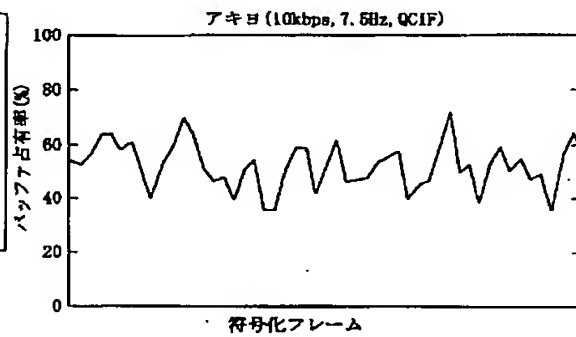
【図8】



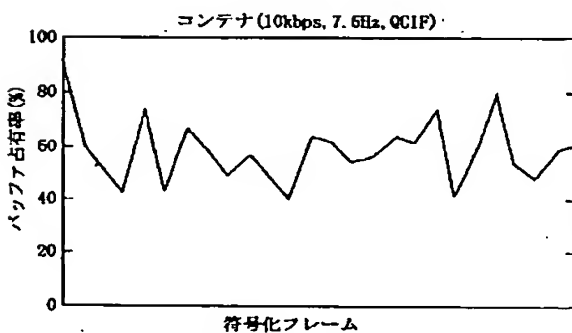
【図3】



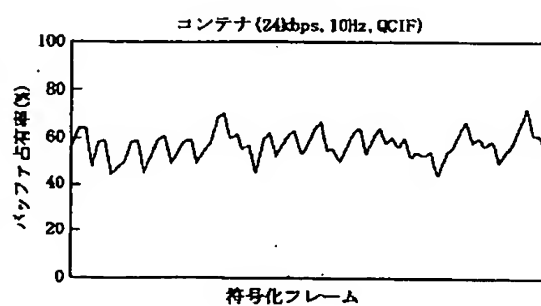
【図6】



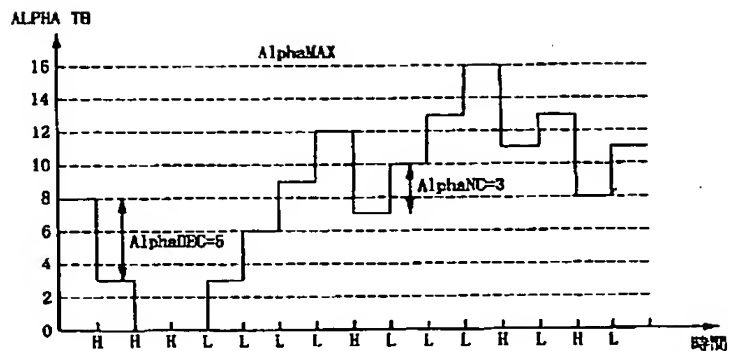
【図7】



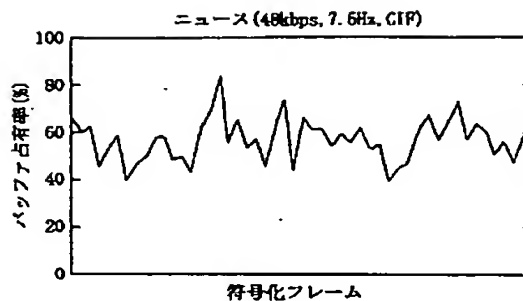
【図9】



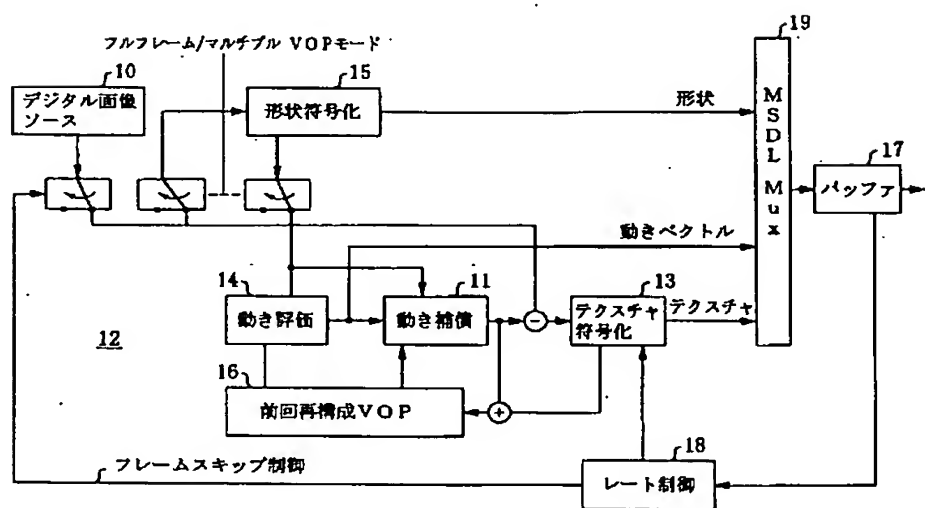
【图 4】



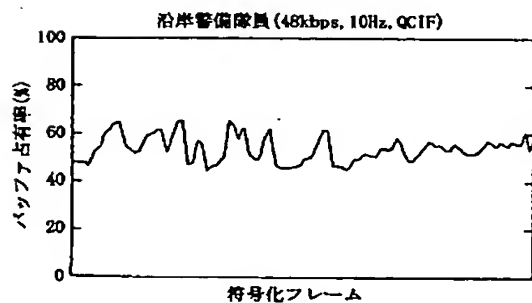
【 10 】



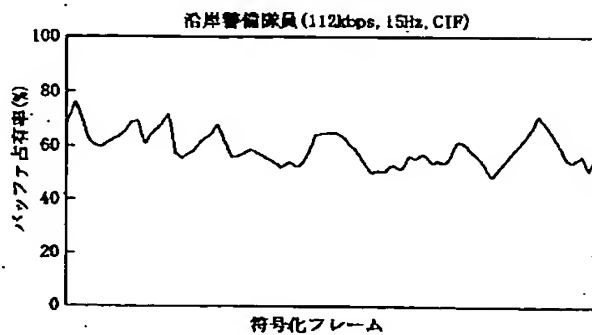
【例 5】



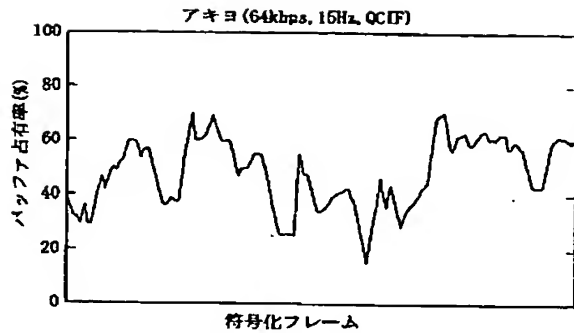
【图 1 1】



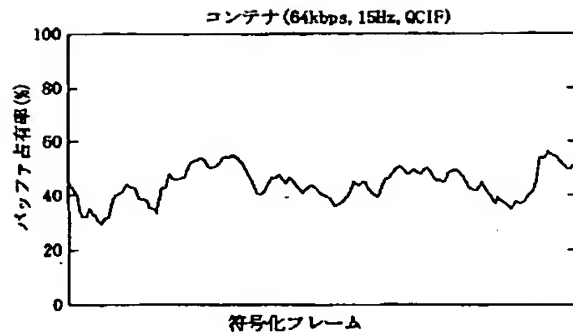
【例 12】



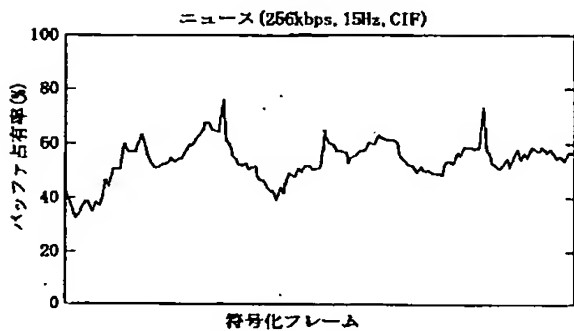
【図 1 3】



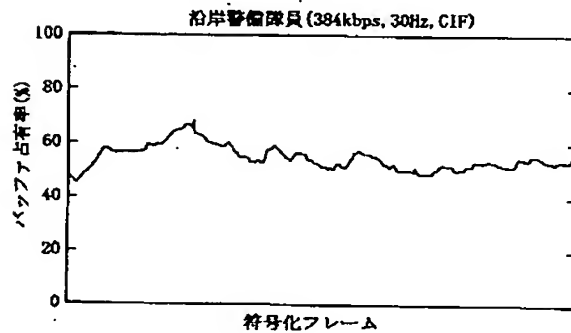
【図 1 4】



【図 1 5】



【図 1 6】



【手続補正書】

【提出日】平成11年10月6日 (1999. 10. 6)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正内容】

【特許請求の範囲】

【請求項1】 可変の入力ビットレートと固定の出力ビットレートとを有するバッファによって圧縮されたビデオ信号を伝送チャネルへ供給するため画像情報のフレームのシーケンスを適応性をもって符号化する方法であって、前記フレームの中の少なくともいくつかは複数のビデオオブジェクトを有し、前記の方法において、変換係数を生成する離散余弦変換を行うためのプロセッサ、及び可変レートで画像表示符号ビットを生成するために前記変換係数を量子化する量子化装置とを含む符号化手段を用いて、フレームセット毎に前記ビデオオブジェクトの各々を符号化し、前記ビデオオブジェクトの各々のテクスチャ、動き及び形状情報を生成する符号化ス

テップと、

前記バッファに前記画像表示符号ビットを格納する格納ステップと、

前記フレームの連続フレーム中の前記ビデオオブジェクトの前記符号化手段によって生成される符号ビットの量を増減するために、二次レートディストーションモデルによる基準値に基づいて前記量子化装置によって利用される量子化パラメータを調整することにより所定のリミット値に基づいて前記バッファの容量を制限する制限ステップと、

ビデオオブジェクトブレン中の一致するビデオオブジェクトに関連する相対的動き、サイズ及び平方偏差パラメータの関数に従って、前記ビデオオブジェクトの間のビデオオブジェクトブレン中の全てのビデオオブジェクトについてビットのターゲット数を配分することにより、第1フレームに続いて所定の時間間隔にわたって発生するシーケンスの中で、各連続フレーム中の各ビデオオブジェクトを符号化するためのビットのターゲット数を算定する算定ステップと、

画質を保ちながら前記バッファのオーバーフローを回避

するために少なくともより高いレートとより低いレートの内の一方で符号化するための前記可変レートを符号化レートとして設定する設定ステップとを有することを特徴とする改良された適応性のあるビデオ符号化方法。

【請求項2】 前記関数がさらに、前記動き、サイズ及び平方偏差パラメータの各々に対する別個の重み係数を有することを特徴とする請求項1に記載の改良された適応性のあるビデオ符号化方法。

【請求項3】 前記平方偏差パラメータが、先行するビデオオブジェクトプレーン中の対応する画素と比較して所定のビデオオブジェクトプレーン中のビデオオブジェクトの各画素に対する平均絶対差値の演算から導出されることを特徴とする請求項1に記載の改良された適応性のあるビデオ符号化方法。

【請求項4】 前記方法がさらに、1つのフレーム中のすべてのビデオオブジェクトの符号化の最後にバッファビット容量と現在のバッファレベルとの間の差が所定のマージンより小さいときに、フレーム周期の間フレームの符号化をスキップするスキップステップを有することを特徴とする請求項1に記載の改良された適応性のあるビデオ符号化方法。

【請求項5】 相対的動き、サイズ及び平方偏差パラメータの前記関数が前記平均絶対差値の2乗に比例する変数を含むことを特徴とする請求項3に記載の改良された適応性のあるビデオ符号化方法。

【請求項6】 前記関数が、

$$T[i] = w_m \times MOT[i] + w_s \times SIZE[i] + w_v \times MAD^2[i]$$

であって、 $MOT[i]$ 、 $SIZE[i]$ 及び $MAD[i]$ は、動き、サイズ及び平均絶対差パラメータの相対比率を示し、 w_m 、 w_s 及び w_v は式

$$w_m + w_s + w_v = 1$$

を満たす重み係数であることを特徴とする請求項3に記載の改良された適応性のあるビデオ符号化方法。

【請求項7】 前記符号化レートが、前記より低いレートと前記より高いレートであるとき、 w_v がそれぞれより低い値とより高い値で選択されることを特徴とする請求項6に記載の改良された適応性のあるビデオ符号化方法。

【請求項8】 前記より低い符号化レートとして

$$w_m = 0.6, w_s = 0.4 \text{ 及び } w_v = 0$$

前記より高い符号化レートとして

$$w_m = 0.25, w_s = 0.25 \text{ 及び } w_v = 0.5$$

として前記重み係数が選択されることを特徴とする請求項7に記載の改良された適応性のあるビデオ符号化方法。

【請求項9】 前記より低い符号化レートとしての前記重み係数が $w_v = 0$ であって、前記より高い符号化レートとして w_v が w_m 又は w_s より大きいことを特徴とする請求項7に記載の改良された適応性のあるビデオ符号化方法。

【請求項10】 前記設定ステップの直前に先行する所

定の時間間隔で連続するスキップされたフレームの数をカウントすることによって前記可変レートの前記設定が決定されることを特徴とする請求項1に記載の改良された適応性のあるビデオ符号化方法。

【請求項11】 前記スキップされたフレーム数が所定の数より小さいとき、前記可変レートが前記より高いレートで設定されることを特徴とする請求項10に記載の改良された適応性のあるビデオ符号化方法。

【請求項12】 前記所定の数が2であることを特徴とする請求項11に記載の改良された適応性のあるビデオ符号化方法。

【請求項13】 前記スキップされたフレーム数が所定の数より小さいとき、前記可変レートが前記より高いレートで設定されることを特徴とする請求項4に記載の改良された適応性のあるビデオ符号化方法。

【請求項14】 可変の入力ビットレートと固定の出力ビットレートとを有するバッファによって圧縮されたビデオ信号を伝送チャネルへ供給するため画像情報のフレームのシーケンスを適応性をもって符号化する方法であって、前記フレームの中の少なくともいくつかは複数のビデオオブジェクトを有し、前記の方法において、変換係数を生成する離散余弦変換を行うためのプロセッサ、及び可変レートで画像表示符号ビットを生成するために前記変換係数を量子化する量子化装置とを含む符号化手段を用いて、フレームセット毎に前記ビデオオブジェクトの各々を符号化し、前記ビデオオブジェクトの各々のテクスチャ、動き及び形状情報を生成する符号化ステップと、

前記バッファに前記画像表示符号ビットを格納する格納ステップと、

前記フレームの連続フレーム中の前記ビデオオブジェクトの前記符号化手段によって生成される符号ビットの量を増減するために、二次レートディストーションモデルによる基準値に基づいて前記量子化装置によって利用される量子化パラメータを調整することにより所定のビット値に基づいて前記バッファの容量を制限する制限ステップと、

ビデオオブジェクトプレーン中の一致するビデオオブジェクトに関連する相対的動き、サイズ及び平方偏差パラメータの関数に従って、前記ビデオオブジェクトの間のビデオオブジェクトプレーン中の全てのビデオオブジェクトについてビットのターゲット数を配分することにより、第1フレームに続いて所定の時間間隔にわたって発生するシーケンスの中で、各連続フレーム中の各ビデオオブジェクトを符号化するステップと、

マスクに従って各ビデオオブジェクトについての前記形状情報を符号化するステップと、

所定の変換率に従って、符号化するための各前記ビデオオブジェクトの各マクロブロックのサイズ変換を行うステップと、

前記各マクロブロックの元のサイズを再構成するステップと、
前記マクロブロックの範囲内の各画素ブロックの変換誤差を決定するステップと、
エラー画素ブロックを識別するために前記変換誤差を所定の閾値と比較するブロックと、
前記閾値を超過しないか最大変換率に到達するまで前記変換率を増やし、変換誤差と前記閾値との比較を再決定するブロックとを有する特徴とする改良された適応性のあるビデオ符号化方法。

【請求項15】 前記方法がさらに、画質を保ちながら前記バッファのオーバーフローを回避するために少なくともより高いレートとより低いレートの内的一方で符号化するための前記可変レートを設定するステップを有することを特徴とする請求項14に記載の改良された適応性のあるビデオ符号化方法。

【請求項16】 前記設定するステップの直前に先行する所定の時間間隔で、連続するスキップされたフレームの数をカウントすることによって前記可変レートの前記設定が決定されることを特徴とする請求項15に記載の改良された適応性のあるビデオ符号化方法。

【請求項17】 相対的動き、サイズ及び平方偏差パラメータの前記関数が前記平均絶対差値の2乗に比例する変数を含むことを特徴とする請求項16に記載の改良された適応性のあるビデオ符号化方法。

【請求項18】 前記関数が、
$$T[i] = W_m \times MOT[i] + W_s \times SIZE[i] + W_v \times MAD^2[i]$$

であって、 $MOT[i]$ 、 $SIZE[i]$ 及び $MAD[i]$ は動き、サイズ及び平均絶対差パラメータの相対比率を示し、 W_m 、 W_s 及び W_v は式

$$W_m + W_s + W_v = 1$$

を満たす重み係数であることを特徴とする請求項17に記載の改良された適応性のあるビデオ符号化方法。

【請求項19】 前記符号化レートが、前記より低いレートと前記より高いレートであるとき、 W_v がそれぞれより低い値とより高い値で選択されることを特徴とする請求項18に記載の改良された適応性のあるビデオ符号化方法。

【請求項20】 可変の入力ビットレートと固定の出力ビットレートとを有するバッファによって圧縮されたビデオ信号を送信チャネルへ供給するため画像情報のフレームのシーケンスを適応性をもって符号化する方法であって、前記フレームの中の少なくともいくつかは複数のビデオオブジェクトを有し、前記の方法において、変換係数を生成する離散余弦変換を行うためのプロセッサ、及び可変レートで画像表示符号ビットを生成するために前記変換係数を量子化する量子化装置とを含む符号化手段を用いて、フレームセット毎に前記ビデオオブジェクトの各々を符号化し、前記ビデオオブジェクトの各々のテクスチャ、動き及び形状情報を生成する符号化ス

テップと、
前記バッファに前記画像表示符号ビットを格納する格納ステップと、

前記フレームの連続フレーム中の前記ビデオオブジェクトの前記符号化手段によって生成される符号ビットの量を増減するために、二次レートディストーションモデルによる基準値に基づいて前記量子化装置によって利用される量子化パラメータを調整することにより所定のリミット値に基づいて前記バッファの容量を制限する制限ステップと、

画質を保ちながら前記バッファのオーバーフローを回避するために少なくともより高いレートとより低いレートの内的一方で符号化するための前記可変レートを設定するステップと、

所定の変換率に従って前記各ビデオオブジェクトの各マクロブロックの前記形状情報のサイズ変換を行うステップと、

前記マクロブロックの範囲内の各画素ブロックの変換誤差を決定するステップと、

エラー画素ブロックを識別するために前記変換誤差を所定のAlpha閾値と比較するステップと、

前記Alpha閾値を超過しないか最大変換率に到達するまで前記変換率を増やし、変換誤差と前記Alpha閾値との比較を再決定するステップと、

1つのフレーム中のすべてのビデオオブジェクトの符号化の最後にバッファビット容量と現在のバッファレベルとの間の差が所定のマージンより小さいときに、フレーム周期の間フレームの符号化をスキップするステップと、

前記設定するステップの直前に先行する所定の時間間隔で、連続するスキップされたフレームの数をカウントすることによって前記可変レートの前記設定が決定され、前記スキップされたフレーム数が所定の数より小さいとき、前記可変レートが前記より高いレートで設定されるステップと、

符号化ステップの後で、前記より高いレート若しくは前記より低いレートの何れが有効であるかを判定するステップと、

前記より低いレートが有効であれば前記Alpha閾値を増やし、前記より高いレートが後続の符号化間隔にとって有効であれば前記Alpha閾値を減らすステップとを有することを特徴とする改良された適応性のあるビデオ符号化方法。

【請求項21】 前記Alpha閾値が初期値としてゼロと最大値との中間の値に設定されることを特徴とする請求項20に記載の改良された適応性のあるビデオ符号化方法。

【請求項22】 前記Alpha閾値が最初の所定のレベルの増分の値で増加し、別の所定のレベルの減分の値で減少することを特徴とする請求項20に記載の改良された

適応性のあるビデオ符号化方法。

【請求項 2 3】 前記増分レベルが前記減分レベルより少ないことを特徴とする請求項 2 2 に記載の改良された適応性のあるビデオ符号化方法。

【請求項 2 4】 前記方法がさらに、ビデオオブジェクトプレーン中の一致するビデオオブジェクトに関連する相対的動き、サイズ及び平方偏差パラメータの関数に従って、前記ビデオオブジェクトの間のビデオオブジェクトプレーン中の全てのビデオオブジェクトについてビットのターゲット数を配分することにより、第 1 フレームに続いて所定の時間間隔にわたって発生するシーケンスの中で、各連続フレーム中の各ビデオオブジェクトを符号化するためのビットのターゲット数を算定する算定ステップを有することを特徴とする請求項 2 1 に記載の改良された適応性のあるビデオ符号化方法。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0 0 1 0

【補正方法】変更

【補正内容】

【0 0 1 0】ISO/IEC JTC1/SC29/WG11 MPEG96/N1469 という名称で、1996年11月にMPEG-4ビデオVM編集に関する専門家班によってそのメンバーに配布された“MPEG-4ビデオ検証モデルバージョン5.0”に解説されている、最新の符号化検証モデル(VM)を実現するための最近の提案で、その内容は本明細書に参考として取り入れられているが、デービッドサーノフ(David Sarnoff)研究センターの代表者達は、“二次レートディストーションモデル(QRD: quadratic rate distortion model)”を用いる新しいレート制御計画”を提案した。「二次レートディストーションモデル(QRD: quadratic rate distortion model)」に関しては、ISO/IEC JTC1/SC29/WG11 MPEG96/N1469という名称で、1996年11月にMPEG-4ビデオVM編集に関する専門家班によってそのメンバーに配布された“MPEG-4ビデオ検証モデルバージョン5.0”に解説されていると共に、資料として同時に提出する Chiang,

T.H.及び Zhang, Y.Q.による「A New Rate Control Scheme Using Quadratic Rate Distortion Model, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 7, No. 1, February 1997, pp.246-250.」に詳しく述べられている。このMPEG-4ビデオ符号化フォーマットは(先行の計画の場合と同様)フレーム毎にエンコーダで可変ビットレートストリームを生成する。この可変ビットレートストリームは固定レートチャンネルによって伝送されなければならないので、チャンネルバッファを用いてビットストリームを円滑に出力する。バッファがオーバーフローまたはアンダーフローするのを回避するために、符号化プロセスのレート制御が必要とされる。

【手続補正 3】

【補正対象書類名】明細書

【補正対象項目名】0 0 1 1

【補正方法】変更

【補正内容】

【0 0 1 1】この最近の提案では、所定のセットのフレーム(ピクチャ)のために符号化プロセスが始まる前に、各フレームのためのターゲットビットレートが計算され、エンコーダの出力からの出力ビットレートが固定ビットレートに抑えられる一方、ピクチャ符号化の結果から生じるビットレートが画像フレームの内容によって(修正されないままになっている場合)比較的広い範囲にわたって変動可能であるという事実を助ける。この提案に従って、各フレームに関連する歪み測度はフレームの平均量子化スケールになるものと仮定されており、歪み測度の逆関数の二次関数としてそのレートディストーション関数がモデル化されている。実際の符号化プロセスが始まる前に、画像のターゲットビットレートは、符号化すべきフレーム数のみならず画像グループを符号化するために残されているビット数によっても評価される。この提案の提唱者達はピクチャレベルでの彼らの計画の実現について触れ、その計画をマクロブロックレベルにまで拡張する可能性についても述べている。

フロントページの続き

(71)出願人 597067574

201 BROADWAY, CAMBRIDGE, MASSACHUSETTS 02139, U. S. A.

(72)発明者 ファイファン・スン

アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ブロードウェイ201 ミツビシ・エレクトリック・インフォメーション・テクノロジー・センター・アメリカ・インコーポレイテッド内

(72)発明者 アンソニー・ヴェトロ

アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ブロードウェイ201 ミツビシ・エレクトリック・インフォメーション・テクノロジー・センター・アメリカ・インコーポレイテッド内

Fターム(参考) 5C059 KK21 KK35 MA00 MA04 MA05
MA14 MA23 MB04 MB10 MB14
MB16 NC11 ME01 ME05 NN08
NN27 PP05 PP06 TA17 TA41
TA60 TB07 TB18 TC00 TC06
TC10 TC25 TD03 TD06 TD12
UA02 UA33